Abstract

For a given sentence grammar, speech recognizers are often required to decode M set of HMMs each of which models a specific acoustic environment. In order to match input acoustic observations to each of the environments, typically recognition search methods require a network of M sub-networks. A new speech recognition search method is described here, which needs only 1 out of the M subnetwork and yet gives the same recognition performance, thus reducing memory requirement for network storage by M-1/M.